Linear Time Series (TS) Models

Financial TS: collection of a financial measurement over time

Example: log return r_t

Data: $\{r_1, r_2, \cdots, r_T\}$ (T data points)

Purpose: What information contained in $\{r_t\}$?

Basic concepts

- Stationarity:
  - Strict: distributions are time-invariant
  - Weak: first 2 moments are time-invariant

  What does weak stationarity mean in practice?

  Past: time plot of $\{r_t\}$ varies around a fixed level within a finite range!

  Future: the first 2 moments of future r_t are the same as those of the data so that meaningful inferences can be made.

- Mean (or expectation) of returns:

$$\mu = E(r_t)$$

- Variance (variability) of returns:

$$\text{Var}(r_t) = E[(r_t - \mu)^2]$$

- Sample mean and sample variance are used to estimate the mean and variance of returns.

$$\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t \quad \& \quad \text{Var}(r_t) = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})^2$$

- Test $H_o : \mu = 0$ vs $H_a : \mu \neq 0$. Compute

$$t = \frac{\bar{r}}{\text{std}(\bar{r})} = \frac{\bar{r}}{\sqrt{\text{Var}(r_t)/T}}$$

  Compare t ratio with $N(0,1)$ dist.

- Lag-k autocovariance:

$$\gamma_k = \text{Cov}(r_t, r_{t-k}) = E[(r_t - \mu)(r_{t-k} - \mu)].$$

- Serial (or auto-) correlations:

$$\rho_\ell = \frac{\text{cov}(r_t, r_{t-\ell})}{\text{var}(r_t)}$$

Note: $\rho_0 = 1$ and $\rho_k = \rho_{-k}$ for $k \neq 0$. Why?

Existence of serial correlations implies that the return is predictable, indicating market inefficience.

- Sample autocorrelation function (ACF)

$$\widehat{\rho}_\ell = \frac{\Sigma_{t=1}^{T-\ell}(r_t - \bar{r})(r_{t+\ell} - \bar{r})}{\Sigma_{t=1}^{T}(r_t - \bar{r})^2},$$

where $\bar{r}$ is the sample mean & T is the sample size.

- Test zero serial correlations (market efficiency)

  – Individual test: for example,

  $H_o : \rho_1 = 0$ vs $H_a : \rho_1 \neq 0$

  $$t = \frac{\hat{\rho}_1}{\sqrt{1/T}} = \sqrt{T}\hat{\rho}_1$$

  Asym. $N(0, 1)$.

  – Joint test (Ljung-Box statistics):

$$H_o : \rho_1 = \cdots = \rho_m = 0 \text{ vs } H_a : \rho_i \neq 0$$

$$Q(m) = T(T+2) \sum_{\ell=1}^{m} \frac{\hat{\rho}_\ell^2}{T-\ell}$$

Asym. chi-squared dist with m degrees of freedom.

- Sources of serial correlations in financial TS
  - Nonsynchronous trading (ch. 5)
  - Bid-ask bounce (ch. 5)
  - Risk premium, etc. (ch. 3)

  Thus, significant sample ACF does not necessarily imply market inefficiency.

Example: Monthly returns of IBM stock from 1926 to 1997.

- R_t: $Q(5) = 5.4(0.37)$ and $Q(10) = 14.1(0.17)$
- r_t: $Q(5) = 5.8(0.33)$ and $Q(10) = 13.7(0.19)$

Remark: What is p-value? How to use it?

Implication: Monthly IBM stock returns do not have significant serial correlations.

Example: Monthly returns of CRSP value-weighted index from 1926 to 1997.

- R_t: $Q(5) = 27.8$ and $Q(10) = 36.0$
- r_t: $Q(5) = 26.9$ and $Q(10) = 32.7$

All highly significant. That is, there exist significant serial correlations in the equal-weighted index returns. (Nonsynchronous trading might explain the existence of the serial correlations, among other reasons.)

Back-shift (lag) operator

A useful notation in TS analysis.

- Definition: $Br_t = r_{t-1}$ or $Lr_t = r_{t-1}$
- $B^2 r_t = B(Br_t) = Br_{t-1} = r_{t-2}$.

B (or L) means time shift! Br_t is the value of the series at time $t-1$.

Suppose that the daily log returns are

Day	1	2	3	4
r_t	0.017	-0.005	-0.014	0.021

Answer the following questions:

- $r_2 =$
- $Br_3 =$
- $B^2 r_5 =$

Question: What is $B2$?

What are the important statistics in practice?

Conditional quantities, not unconditional

A proper perspective: at a time point t

- Available data: $\{r_1, r_2, \cdots, r_{t-1}\} \equiv F_{t-1}$
- The return is decomposed into two parts as

$$\begin{aligned} r_t &= \text{predictiable part} + \text{not predic. part} \\ &= \text{function of elements of } F_{t-1} + a_t \end{aligned}$$

In other words, given information F_{t-1}

$$\begin{aligned} r_t &= \mu_t + a_t \\ &= E(r_t|F_{t-1}) + \sigma_t \epsilon_t \end{aligned}$$

– μ_t: conditional mean of r_t

– a_t: shock or innovation at time t

– ϵ_t: an iid sequence with mean zero and variance 1

– σ_t: conditional standard deviation (commonly called volatility in finance)

Traditional TS modeling is concerned with μ_t:

Model for μ_t: **mean equation**

Volatility modeling concerns σ_t.

Model for σ_t^2: **volatility equation**

Univariate TS analysis serves two purposes

- a model for μ_t
- understanding models for σ_t^2: properties, forecasting, etc.

Linear time series: r_t is linear if

- the predictiable part is a linear function of F_{t-1}
- $\{a_t\}$ are indep. and have the same dist. (iid)

Mathematically, it means r_t can be written as

$$r_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i},$$

where μ is a constant, $\psi_0 = 1$ and $\{a_t\}$ is an iid sequence with mean zero and well-defined distribution.

In the economic literature, a_t is the *shock* (or *innovation*) at time t and $\{\psi_i\}$ are the *impulse* responses of r_t.

White noise: iid sequence (with finite variance), which is the building block of linear TS models.

White noise is not predictable, but has zero mean and finite variance.

Univariate linear time series models

1. autoregressive (AR) models

2. moving-average (MA) models

3. mixed ARMA models

4. seasonal models

5. regression models with time series errors

6. fractionally differenced models (long-memory)

Example Quarterly growth rate of U.S. real gross national product (GNP), seasonally adjusted, from the second quarter of 1947 to the first quarter of 1991.

An AR(3) model for the data is

$$r_t = 0.005 + 0.35 r_{t-1} + 0.18 r_{t-2} - 0.14 r_{t-3} + a_t, \quad \hat{\sigma}_a = 0.01,$$

where $\{a_t\}$ denotes a white noise with variance σ_a^2. Given r_n, r_{n-1} & r_{n-2}, we can predict r_{n+1} as

$$\hat{r}_{n+1} = 0.005 + 0.35 r_n + 0.18 r_{n-1} - 0.14 r_{n-2}.$$

Other implications of the model?

Example: Monthly simple return of CRSP equal-weighted index

$$R_t = 0.013 + a_t + 0.178a_{t-1} - 0.13a_{t-3} + 0.135a_{t-9}, \quad \hat{\sigma}_a = 0.073$$

Checking: $Q(10) = 11.4(0.122)$ for the residual series a_t.

Implications of the model?

Important properties of a model

- Stationarity condition
- Basic properties: mean, variance, serial dependence
- Empirical model building: specification, estimation, & checking
- Forecasting

Simple AR models (regression ?)

AR(1) model:

1. Form: $r_t = \phi_0 + \phi_1 r_{t-1} + a_t$, where ϕ_0 and ϕ_1 are real numbers, which are referred to as "parameters" (to

be estimated from the data in an application). For example,

$$r_t = 0.005 + 0.2r_{t-1} + a_t$$

2. Stationarity: necessary and sufficient condition $|\phi_1| < 1$. Why?

3. Mean: $E(r_t) = \frac{\phi_0}{1-\phi_1}$

4. Variance: $\text{Var}(r_t) = \frac{\sigma_a^2}{1-\phi_1^2}$.

5. Autocorrelations: $\rho_1 = \phi_1, \rho_2 = \phi_1^2$, etc. In general, $\rho_k = \phi_1^k$ and ACF ρ_k decays exponentially as k increases,

6. Forecast (minimum squared error):

   (a) 1-step ahead forecast at time n, the forecast origin:

   $$\hat{r}_n(1) = \phi_0 + \phi_1 r_n$$

   (b) 1-step ahead forecast error:

   $$e_n(1) = r_{n+1} - \hat{r}_n(1) = a_{n+1}$$

Thus, a_{n+1} is the *un-predictable* part of r_{n+1}. It is the shock at time $n+1$!

(c) Variance of 1-step ahead forecast error:

$$\text{Var}[e_n(1)] = \text{Var}(a_{n+1}) = \sigma_a^2.$$

(d) 2-step ahead forecast:

$$\hat{r}_n(2) = \phi_0 + \phi_1 \hat{r}_n(1)$$

(e) 2-step ahead forecast error:

$$e_n(2) = r_{n+2} - \hat{r}_n(2) = a_{n+2} + \phi_1 a_{n+1}$$

(f) Variance of 2-step ahead forecast error:

$$\text{Var}[e_n(2)] = (1 + \phi_1^2)\sigma_a^2$$

which is greater than or equal to $\text{Var}[e_n(1)]$, implying that uncertainty in forecasts increases as the number of steps increases.

(g) Behavior of multi-step ahead forecasts.

7. A compact form: $(1 - \phi_1 B) r_t = \phi_0 + a_t$.

AR(2) model:

1. Form: $r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + a_t$, or

$$(1 - \phi_1 B - \phi_2 B^2) r_t = \phi_0 + a_t.$$

2. Stationarity condition: (factor of polynomial)

3. Characteristic equation: $(1 - \phi_1 x - \phi_2 x^2) = 0$

4. Mean: $E(r_t) = \frac{\phi_0}{1-\phi_1-\phi_2}$

5. ACF: $\rho_0 = 1$, $\rho_1 = \frac{\phi_1}{1-\phi_2}$,

$$\rho_\ell = \phi_1 \rho_{\ell-1} + \phi_2 \rho_{\ell-1}, \quad \ell \geq 2.$$

6. Stochastic business cycle: if $\phi_1^2 + 4\phi_2 < 0$, then r_t shows characteristics of business cycles with average length

$$k = \frac{360^o}{\cos^{-1}[\phi_1/(2\sqrt{-\phi_2})]}.$$

7. Forecasts: Similar to AR(1) models

Example: US GNP growth rate series revisited.

Building an AR model

- Order specification

  1. Partial ACF: (naive, but effective)
     - Use consecutive fittings
     - See Text (p. 36) for details
     - Key feature: PACF cuts off at lag p for an AR(p) model.
     - Illustration
  2. Akaike information criterion

     Find the AR order with *minimum* AIC.

- Needs a constant term? Check the sample mean.

- Estimation: least squares method or maximum likelihood method

- Model checking:

  1. Residual: obs minus the fit, i.e. 1-step ahead forecast errors at each time point.

2. Residual should be close to white noise if the model is adequate. Use Ljung-Box statistics of residuals, but degrees of freedom is $m - g$, where g is the number of AR coefficients used in the model.

- Many software packages available, e.g. SCA, SAS, SPSS, etc.

Example: See section 2.4 of the text.

Moving-average (MA) model

Model with finite time lags of memory!

Some daily stock returns have minor serial correlations.

Can be modeled as MA or AR models.

MA(1) model

- Form: $r_t = \mu + a_t - \theta a_{t-1}$
- Stationarity: always stationary.
- Mean (or expectation): $E(r_t) = \mu$
- Variance: $\text{Var}(r_t) = (1 + \theta^2)\sigma_a^2$.

- Autocovariance:

  1. Lag 1: $\text{Cov}(r_t, r_{t-1}) = -\theta\sigma_a^2$
  2. Lag ℓ: $\text{Cov}(r_t, r_{t-\ell}) = 0$ for $\ell > 1$.

  Thus, r_t is not related to $r_{t-2}, r_{t-3}, \cdots$.

- ACF: $\rho_1 = \frac{-\theta}{1+\theta^2}$, $\rho_\ell = 0$ for $\ell > 1$.

  Finite memory! MA(1) models do not remember what happen two time periods ago.

- Forecast (at origin $t = n$):

  1. 1-step ahead: $\hat{r}_n(1) = \mu - \theta a_n$. Why? Because at time n, a_n is known, but a_{n+1} is not.
  2. 1-step ahead forecast error: $e_n(1) = a_{n+1}$ with variance σ_a^2.
  3. Multi-step ahead: $\hat{r}_n(\ell) = \mu$ for $\ell > 1$.

     Thus, for an MA(1) model, the multi-step ahead forecasts are just the mean of the series. Why? Because the model has memory of 1 time period.

4. Multi-step ahead forecast error:

$$e_n(\ell) = a_{n+\ell} - \theta a_{n+\ell-1}$$

5. Variance of multi-step ahead forecast error:

   $(1+\theta^2)\sigma_a^2 =$ variance of r_t.

- Invertibility:
  - Concept: r_t is a proper linear combination of a_t and the past observations $\{r_{t-1}, r_{t-2}, \cdots\}$.
  - Why is it important? It provides a simple way to obtain the shock a_t.

    For an invertible model, the dependence of r_t on $r_{t-\ell}$ converges to zero as ℓ increases.
  - Condition: $|\theta| < 1$.
  - Invertibility of MA models is the dual property of stationarity for AR models.

<u>MA(2) model</u>

- Form: $r_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$. or

$$r_t = \mu + (1 - \theta_1 B - \theta_2 B^2) a_t.$$

- Stationary with $E(r_t) = \mu$.
- Variance: $\text{Var}(r_t) = (1 + \theta_1^2 + \theta_2^2)\sigma_a^2$.
- ACF: $\rho_2 \neq 0$,but $\rho_\ell = 0$ for $\ell > 2$.
- Forecasts go the the mean after 2 periods.

Building an MA model

- Specification: Use sample ACF

  Sample ACFs are all small after lag q for an MA(q) series. (See test of ACF.)
- Constant term? Check the sample mean.
- Estimation: use maximum likelihood method
  - Conditional: Assume $a_t = 0$ for $t \leq 0$

– Exact: Treat a_t with $t \leq 0$ as parameters, estimate them to obtain the likelihood function.

Exact method is preferred, but it is more computing intensive.

- Model checking: examine residuals (to be white noise)
- Forecast: use the residuals as $\{a_t\}$ (which can be obtained from the data and fitted parameters) to perform forecasts.

Example: see the text.

Mixed ARMA model: A compact form for flexible models.

Focus on the ARMA(1,1) model for

1. simplicity
2. useful for understanding GARCH models in Ch. 3 for volatility modeling.

ARMA(1,1) model

- Form: $(1-\phi_1 B)r_t = \phi_0 + (1-\theta B)a_t$ or

$$r_t = \phi_1 r_{t-1} + \phi_0 + a_t - \theta_1 a_{t-1}.$$

  A combination of an AR(1) on the LHS and an MA(1) on the RHS.

- Stationarity: same as AR(1)
- Invertibility: same as MA(1)
- Mean: as AR(1), i.e. $E(r_t) = \frac{\phi_0}{1-\phi_1}$
- Variance: given in the text
- ACF: Satisfies $\rho_k = \phi_1 \rho_{k-1}$ for $k > 1$, but

$$\rho_1 = \phi_1 - [\theta_1 \sigma_a^2 / \mathrm{Var}(r_t)] \neq \phi_1.$$

  This is the difference between AR(1) and ARMA(1,1) models.

- PACF: does not cut off at finite lags.

Building an ARMA(1,1) model

- Specification: use EACF or AIC
- What is EACF? How to use it? [See text].
- Estimation: cond. or exact likelihood method
- Model checking: as before
- Forecast: MA(1) affects the 1-step ahead forecast. Others are similar to those of AR(1) models.

Three model representations:

- ARMA form: compact, useful in estimation and forecasting
- AR representation: (by long division)

$$r_t = \phi_0 + a_t + \pi_1 r_{t-1} + \pi_2 r_{t-2} + \cdots$$

  It tells how r_t depends on its past values.

- MA representation: (by long division)

$$r_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots$$

It tells how r_t depends on the past shocks.

For a stationary series, ψ_i converges to zero as $i \to \infty$. Thus, the effect of any shock is transitory.

The MA representation is particularly useful in computing variances of forecast errors.

For a ℓ-step ahead forecast, the forecast error is

$$e_n(\ell) = a_{n+\ell} + \psi_1 a_{n+\ell-1} + \cdots + \psi_{\ell-1} a_{n+1}.$$

The variance of forecast error is

$$\mathrm{Var}[e_n(\ell)] = (1 + \psi_1^2 + \cdots + \psi_{\ell-1}^2)\sigma_a^2.$$

Unit-root Nonstationarity

<u>Random walk</u>

- Form $p_t = p_{t-1} + a_t$
- Unit root? It is an AR(1) model with coefficient $\phi_1 = 1$.
- Nonstationary: Why? Because the variance of r_t diverges to infinity as t increases.

- Strong memory: sample ACF approaches 1 for any finite lag.

- Repeated substitution shows

$$p_t = \sum_{i=0}^{\infty} a_{t-i} = \sum_{i=0}^{\infty} \psi_i a_{t-i}$$

where $\psi_i = 1$ for all i. Thus, ψ_i does not converge to zero. The effect of any shock is permanent.

<u>Random walk with a drift</u>

- Form: $p_t = \mu + p_{t-1} + a_t, \quad \mu \neq 0.$

- Has a unit root

- Nonstationary

- Strong memory

- Has a time trend with slope μ. Why?

<u>differencing</u>

- 1st difference: $r_t = p_t - p_{t-1}$

If p_t is the log price, then the 1st difference is simply the log return. Typically, 1st difference means the "change" or "increment" of the original series.

- Seasonal difference: $y_t = p_t - p_{t-s}$, where s is the periodicity, e.g. $s = 4$ for quarterly series and $s = 12$ for monthly series.

  If p_t denotes quarterly earnings, then y_t is the change in earning from the same quarter one year before.

Meaning of the constant term in a model

- MA model: mean
- AR model: related to mean
- 1st differenced: time slope, etc.

Practical implication in financial time series

Example: Monthly log returns of General Electrics (GE) from 1926 to 1999 (74 years)

Sample mean: 1.04%, $\text{std}(\hat{\mu}) = 0.26$

Very significant!

is about 12.45% a year

$1 investment in the beginning of 1926 is worth

- annual compounded payment: $5907
- quarterly compounded payment: $8720
- monthly compounded payment: $9570
- Continuously compounded?

Seasonal Time Series

- TS with periodic pattern, e.g. quarterly earnings
- Useful in weather-related derivative pricing
- Useful in analysis of transactions data (high-frequency data), e.g. U-shaped pattern in intraday data

Example Quarterly earnings of Johnson & Johnson

See the time plot and sample ACFs

Multiplicative model

Airline model (for quarterly series)

- Form:

$$r_t - r_{t-1} - r_{t-4} + r_{t-5} = a_t - \theta_1 a_{t-1} - \theta_4 a_{t-4} + \theta_1\theta_4 a_{t-5}$$

or

$$(1-B)(1-B^4)r_t = (1-\theta_1 B)(1-\theta_4 B^4)a_t$$

- Define the differenced series w_t as

$$w_t = r_t - r_{t-1} - r_{t-4} + r_{t-5} = (r_t - r_{t-1}) - (r_{t-4} - r_{t-5}).$$

It is called *regular* and *seasonal* differenced series.

- ACF of w_t has a nice symmetric structure (see the text), i.e. $\rho_{s-1} = \rho_{s+1} = \rho_1\rho_s$. Also, $\rho_\ell = 0$ for $\ell > s+1$.

- This model is widely applicable to many many seasonal time series.

- Multiplicative model means that the regular and seasonal dependences are roughly orthogonal to each other.
- Forecasts: exhibit same pattern as the observed series.

Example Detailed analysis of J&J earnings.

Regression Models with Time Series Errors

- Has many applications
- Impact of serial correlations in regression is often overlooked.

  It may introduce biases in estimates and in standard errors, resulting in unreliable t-ratios.
- Detecting residual serial correlation: Use Q-stat instead of DW-statistic, which is not sufficient!
- Joint estimation of all parameters is preferred; see SCA.
- Proper analysis: via illustration

Example. U.S. interest rate data

Long-memory models

- Meaning? ACF decays to zero very slowly!
- Example: ACF of squared or absolute log returns

  ACFs are small, but decay very slowly.
- How to model long memory? Use "fractional" difference: namely, $(1 - B)^d r_t$, where $-0.5 < d < 0.5$.
- Importance? In theory, Yes. In practice, yet to be determined.

Summary of the chapter

- Sample ACF $\Rightarrow$ MA order
- Sample PACF $\Rightarrow$ AR order
- A simple model for "conditional mean" of a FTS can be obtained in SCA via "iarima" with the help of ACF, PACF and EACF
- Check a fitted model before forecasting, e.g. residual ACF and hetroschedasticity (chapter 3)
- Interpretation of a model, e.g. constant term &

  For an AR(1) with coefficient ϕ_1, the speed of mean reverting as measured by half-life is

  $$k = \frac{\ln(0.5)}{\ln(\phi_1)}.$$

  For an MA model, in a finite number of periods.
- Make practical use of regression models with time series errors, e.g. regression with AR(1) residuals

Perform a joint estimation instead of two-step procedures, e.g. Cochrane-Orcutt (1949).

Example: Is there a Friday effect on asset returns?

If a daily market index is used, serial correlation may exist.

See Exercise 8 of Chapter 2.

- Basic properties of a random-walk model
- Multiplicative seasonal models, especially the so-called airline model.